



Characterization of Aquilegia Polycomb Repressive Complex 2 homologs reveals absence of imprinting

Citation

Gleason, Emily J., and Elena M. Kramer. 2012. "Characterization of Aquilegia Polycomb Repressive Complex 2 Homologs Reveals Absence of Imprinting." *Gene* 507 (1) (October): 54–60.

Published Version

doi:10.1016/j.gene.2012.07.004

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:11828627>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

Characterization of *Aquilegia* Polycomb Repressive Complex 2 homologs reveals absence of imprinting

Emily J. Gleason^{a,b} and Elena M. Kramer^{a*}

^aDept. of Organismic and Evolutionary Biology, Harvard University, 16 Divinity Ave.
Cambridge, MA, 02138, USA. ekramer@oeb.harvard.edu

^bDept. of Molecular and Cellular Biology, Harvard University, 16 Divinity Ave.
Cambridge, MA, 02138, USA. egleason@fas.harvard.edu

* Author for correspondence: Dept. of Organismic and Evolutionary Biology, Harvard University, 16 Divinity Ave. Cambridge, MA, 02138, USA. (office) 617-496-3460, (fax) 617-496-5854, ekramer@oeb.harvard.edu.

Abstract

Epigenetic regulation is important for maintaining gene expression patterns in multicellular organisms. The Polycomb Group (PcG) proteins form several complexes with important and deeply conserved epigenetic functions in both the plant and animal kingdoms. The plant Polycomb Repressive Complex 2 (PRC2) contains four core proteins, Enhancer of Zeste (E(z)), Suppressor of Zeste 12 (Su(z)12), Extra Sex Combs (ESC), and Multi-Copy Suppressor of IRA 1 (MSI1), and functions in many developmental transitions. In some plant species, including rice and Arabidopsis, duplications in the core PRC2 proteins allow the formation of PRC2s with distinct developmental functions. In addition, members of the plant specific VEL PHD family have been shown to associate with the PRC2 complex in Arabidopsis and may play a role in targeting the PRC2 to specific loci. Here we examine the evolution and expression of the PRC2 and VEL PHD families in *Aquilegia*, a member of the lower eudicot order Ranunculales and an emerging model for the investigation of plant ecology, evolution and developmental genetics. We find that *Aquilegia* has a relatively simple PRC2 with only one homolog of *Su(z)12*, *ESC* and *MSI1* and two ancient copies of *E(z)*, *AqSWN* and *AqCLF*. *Aquilegia* has four members of the VEL PHD family, three of which appear to be closely related to Arabidopsis proteins known to associate with the PRC2. The PRC2 and VEL PHD family proteins are expressed at a relatively constant level throughout *A. vulgaris* development, with the VEL PHD family and *MSI1* expressed at higher levels during and after vernalization and in the inflorescence. Both *AqSWN* and *AqCLF* are expressed in *Aquilegia* endosperm but neither copy is imprinted.

Highlights

- *Aquilegia* has a simple complement of PRC2 homologs with no recent duplications.
- The PRC2 genes are broadly expressed throughout *A. vulgaris* development.
- The ancient paralogs *AqCLF* and *AqSWN* are not imprinted in *Aquilegia* endosperm.
- VEL PHD gene expression in *A. vulgaris* is not confined to vernalization.

Keywords

Polycomb Repressive Complex 2 (PRC2), VERNALIZATION5/VIN3-Like Plant Homeodomain (VEL PHD) Family, Imprinting, Evolution, *Aquilegia*.

Abbreviations

<i>Aq</i> – <i>Aquilegia</i>	DFCI – Dana-Farber Cancer Institute
<i>AqC</i> – <i>Aquilegia canadensis</i>	DNA – Deoxyribonucleic Acid
<i>AqV</i> – <i>Aquilegia vulgaris</i>	DNase – deoxyribonuclease
<i>At</i> – <i>Arabidopsis thaliana</i>	<i>E(z)</i> – <i>Enhancer of Zeste</i>
BLAST – Basic Local Alignment Search Tool	Ed – Endosperm
bp – base pair(s)	Em – Embryo
CAF1 – Chromatin Assembly Factor 1	<i>EMF2</i> – <i>EMBRYONIC FLOWER 2</i>
cDNA - DNA complementary to RNA	<i>ESC</i> – <i>Extra Sex Combs</i>
ChIP – Chromatin Immuno-Precipitation	EST – Expressed Sequence Tag
<i>CLF</i> – <i>CURLY LEAF</i>	Eudicots – eudicotyledonous

FIE – FERTILIZATION INDEPENDENT

ENDOSPERM

FIS2 – FERTILIZATION

INDEPENDENT SEED 2

FLC – FLOWERING LOCUS C

H3K27 – Histone H3 Lysine 27

HOX – Homeobox

Inflo – Inflorescence

IPP2 - Isopentyl

Pyrophosphate:Dimethylallyl

Pyrophosphate Isomerase

Mbp – Mega base pair

MEA – MEDEA

Mez1 – Maize E(z)1

ML – Maximum Likelihood

MSI1 – MULTI COPY SUPPRESSOR

OF IRA 1

MSI2 - MULTI COPY SUPPRESSOR

OF IRA 2

MSI3 - MULTI COPY SUPPRESSOR

OF IRA 3

MSI5 – MULTI COPY SUPPRESSOR

OF IRA 5

n – chromosome number

oligo – oligodeoxyribonucleotide

Os – Oryza sativa

PcG – Polycomb Group

PCR – Polymerase Chain Reaction

Pp – Physcomitrella patens

PRC1 –Polycomb Repressive Complex1

PRC2 –Polycomb Repressive Complex2

qRT-PCR – quantitative Real Time

Polymerase Chain Reaction

RACE - Rapid Amplification of cDNA

Ends

RAxML- Randomized Axelerated

Maximum Likelihood

RNA – Ribonucleic Acid

RT-PCR - Reverse Transcriptase PCR

SNP – Single Nucleotide Polymorphism

Su(z)12 – Suppressor of Zeste12

SWN – SWINGER

UTR – Untranslated region(s)

VEL 1 – VERNALIZATION5/VIN3-LIKE

PROTEIN 1

VEL 2 – VERNALIZATION5/VIN3-LIKE

PROTEIN 2
VEL 3 – VERNALIZATION5/VIN3-LIKE

PROTEIN 3
VEL PHD – Vernalization 5/VIN3-Like
Plant Homeodomain

VIN3 – VERNALIZATION

INSENSITIVE3

VRN1 – VERNALIZATION 1

VRN5 – VERNALIZATION 5

Zm – Zea mays

Database ID: JN944598- JN944605

Section 1: Introduction

The last common ancestor of plants and animals lived approximately 1.6 billion years ago, before the evolution of multicellular organisms (reviewed in Meyerowitz, 2002). Thus, multicellularity most likely arose independently in these two groups and, accordingly, many aspects of their development are very different. However, in both lineages the maintenance of proper gene expression in differentiated cells is essential for the development of multicellularity. Gene expression is maintained in a heritable fashion via a process of cellular memory known as epigenetics (Holliday, 1994; Russo et al., 1996; Feil, 2008). Many proteins involved in epigenetic maintenance of gene expression are highly conserved between plants and animals and appear to function in a remarkably similar manner (Pien and Grossniklaus, 2007; Whitcomb et al., 2007; Köhler and Hennig, 2010).

One key example is the Polycomb Group (PcG), a set of proteins with important and deeply conserved functions in epigenetic silencing. These proteins were first discovered in *Drosophila melanogaster* as repressors of the HOX genes (Lewis, 1978). Although several PcG complexes exist in both plants and animals, each with distinct functions in epigenetic silencing, only the Polycomb Repressive Complex 2 (PRC2) is thought to be conserved between plants and animals (Schuettengruber et al., 2007; Whitcomb et al., 2007). Recently a complex has been identified in Arabidopsis that may have PRC1-like function (Xu and Shen, 2008; Bratzel et al., 2010), but this complex appears to include RING finger proteins, similar to the animal PRC1 complex, as well as both LHP1, a plant homolog of the animal protein HP1 that is not found in the animal PRC1 complex,

and EMF1, a plant specific protein (Calonje et al., 2008; Xu and Shen, 2008; Exner et al., 2009; Bratzel et al., 2010; Beh et al., 2012). Thus, it appears that while the PRC2 complex members are genetically homologous between multicellular organisms, the plant protein complex that plays a functionally analogous role to PRC1 is largely composed of subunits that are not homologous to members of the animal PRC1 complex.

The main function of the PRC2 complex appears to be trimethylation of lysine 27 of histone H3 (H3K27), a histone modification known to suppress gene expression (Schubert et al., 2006). The PRC2 contains four core proteins; the histone methyltransferase *Enhancer of Zeste (E(z))* and three other proteins thought to enhance PRC2 binding to nucleosome; *Suppressor of Zeste 12 (Su(z)12)*, *Extra Sex Combs (ESC)*, and *Multi-Copy Suppressor of IRA 1 (MSI1)* (Nekrasov et al., 2005; Pien and Grossniklaus, 2007). In some plant species, including rice and Arabidopsis, duplications in the core PRC2 proteins allow these species to form PRC2's with distinct developmental functions (Whitcomb et al., 2007; Luo et al., 2009). Recent studies have shown that the PRC2 is involved in developmental transitions in a number of plant species. In the plant model system *Arabidopsis thaliana*, PRC2s function in many processes including endosperm development, early repression of flowering to allow proper vegetative development, the eventual transition to flowering, and flower organogenesis (Goodrich et al., 1997; Gendall et al., 2001; Yoshida et al., 2001; Kohler et al., 2003). In rice, the mutant phenotype of *OsEMF2b*, suggests that the PRC2 complex may play a role in floral induction under long days, flower development, and

suppressing cell divisions in the unfertilized ovule in rice (Luo et al., 2009). Likewise, ChIP analysis of the barley floral promotion locus *VERNALIZATION 1* (*VRN1*) before and after vernalization showed that regulatory regions contained differential levels of H3K27 trimethylation, the histone modification deposited by the PRC2 complex (Oliver et al., 2009). This suggests that the PRC2 complex may function in floral induction in barley as well (Oliver et al., 2009). In the moss species, *Physcomitrella patens*, deletion of the PRC2 genes *PpCLF* and *PpFIE* induces sporophyte-like development and gene expression in the gametophyte, indicating that PRC2-dependent remodeling may be required for the switch from gametophyte to sporophyte development (Mosquna et al., 2009; Okano et al., 2009).

Consistent with these common roles in regulating life stages and tissue identity, another component of PRC2 function in flowering plants is a role in differential imprinting of loci in the maternal and paternal genomes of developing embryos and endosperm, the latter being a nutritive tissue containing two maternal and one paternal genomic complements (Baroux et al., 2002). Furthermore, members of the PRC2 complex itself have been found to be imprinted in *Arabidopsis* and several grasses (Kinoshita et al., 1999; Springer et al., 2002; Guitton et al., 2004; Luo et al., 2009). In the cases of *MEDEA* (*MEA*), *FERTILIZATION INDEPENDENT SEED 2* (*FIS2*), and *FERTILIZATION INDEPENDENT ENDOSPERM* (*FIE*), the loci are imprinted in the endosperm such that the maternal copies are expressed while the paternal copies are silenced (Guitton et al., 2004). Recent work has demonstrated that a maize E(z)-like gene (*Mez1*), maize *ZmFIE1*, and rice *OsFIE1* are similarly imprinted in the endosperm, suggesting that PcG

imprinting may be a common theme in endosperm development (Springer et al., 2002; Haun et al., 2007; Luo et al., 2009; Rodrigues et al., 2010).

In both plants and animals, PRC2s are thought to associate with other proteins that help recruit them to specific loci (Köhler and Hennig, 2010; Margueron and Reinberg, 2011). In *Arabidopsis*, members of a plant specific group known as the VIL (VIN3-like) or VEL PHD family have been shown to associate with the PRC2 complex and seem to be required for PRC2 repression of the floral repressor *FLC* during and after vernalization (Sung et al., 2006; Greb et al., 2007; De Lucia et al., 2008). Intriguingly, VEL PHD homologs are also induced by vernalization in wheat, despite the fact that the grasses evolved their cold response independently (Fu et al., 2007). It remains to be determined, however, whether these wheat genes are actually functioning in the floral promotion pathway.

Here we examine the evolution and expression of the PRC2 and VEL PHD families in the emerging model system, *Aquilegia*. The genus *Aquilegia* has been the subject of ecological, evolutionary and genetic studies for over 50 years (reviewed in Hodges and Kramer, 2007). *Aquilegia* is of interest for a number of reasons. First, *Aquilegia* has a small genome ($n=7$, approximately 300 Mbp) with a number of genetic and genomic tools, including an extensive EST database and the recently sequenced *Aquilegia coerulea* genome (reviewed in Kramer, 2009). Second, as a member of the order Ranunculales, an early diverging lineage of the eudicotyledonous flowering plants that arose before the radiation of the core eudicots, it represents a rough phylogenetic

midpoint between *Arabidopsis* and model systems in the grasses (reviewed in Kramer and Hodges, 2010). Additionally *Aquilegia* has a number of interesting morphological and physiological features including vernalization-based control of flowering which is thought to represent what is likely to be an independent derivation of vernalization response relative to *Arabidopsis* and the grasses (Ballerini and Kramer, 2011). Finally, *Aquilegia* has undergone a recent adaptive radiation, resulting in low sequence variation and a high degree of fertility between species. This allows the use of multiple different species as models as well as the use of interspecific crosses to test phenomena such as imprinting.

In the current study, we have performed broad identification of chromatin remodeling homologs in the recently sequenced *A. coerulea* genome with more detailed study of PRC2 and VEL PHD homologs. The strongly vernalization-responsive species *A. vulgaris* was further utilized to determine broad expression patterns over a range of tissue types and developmental stages. Lastly, we used interspecific crosses and naturally occurring polymorphism to investigate patterns of imprinting in the paralogous *AqCLF* and *AqSWN* loci. This work lays the foundation for future studies of epigenetic modification in the lower eudicots model *Aquilegia* and provides sequence data for broadly evolutionary studies of numerous gene families.

Section 2: Materials and Methods

2.1 Gene cloning

In order to identify genes of interest in the *Aquilegia* genome, BLAST searches (Altschul et al., 1990) of the *Aquilegia* DFCI Gene Index (<http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=Aquilegia>) and the *Aquilegia coerulea* genome (http://www.phytozome.net/search.php?method=Org_Acoerulea) were performed using the sequences of our genes of interest from *Arabidopsis thaliana* or, in a few cases, from *Vitis*.

In the cases of *AqFIE*, *AqEMF2*, and *AqCLF*, BLAST searches did not identify the full length sequence, so 3' and 5' Rapid Amplification of cDNA Ends (RACE) was used to determine the complete sequence. The targeted loci were amplified from a mix of cDNA prepared from RNA isolated from young leaves and primers designed based on the fragments obtained above (see Supplemental Table 1 for primer sequences). 5' RACE followed the 5' RACE System for Rapid Amplification of cDNA Ends, Version 2.0 protocol (Invitrogen, Carlsbad, CA). 3' RACE was performed as described in Kramer et al. (2003). Fragments were cloned using the TOPO-TA Cloning Kit and TOP10 competent cells (Invitrogen, Carlsbad, CA) and several clones per cloning reaction were sequenced using Big Dye v3.1 (Life Technologies Corporation, Carlsbad, CA).

In the case of *AqSWN*, *AqVIN3A*, *AqVIN3B*, and *AqVRN5*, BLAST searches did not identify an EST or predicted an open reading frame, so we used a BLAST search of the *Aquilegia coerulea* genome

(http://www.phytozome.net/search.php?method=Org_Acoerulea) to identify regions that showed similarity to the query sequence. We then used the Soft Berry FGENESH program

(<http://linux1.softberry.com/berry.phtml?topic=fgenes&group=programs&subgroup=gfind>) to predict open reading frames for the loci. cDNA sequences were confirmed using specific primers designed for internal Reverse Transcriptase PCR (RT-PCR) as described in Kramer et al. (2003) as well as 5' RACE for *AqVIN3B* and 3' RACE for *AqSWN* as described above. All new sequences are deposited in Genbank under accession numbers JN944598- JN944605.

2.2 Phylogenetic analysis

For all gene trees, homologs to the PRC2 genes and VEL PHD family were identified for a variety of land plant taxa by using the BLAST algorithm to search GenBank, the DFCI Plant Gene Index (<http://compbio.dfci.harvard.edu/tgi/plant.html>), the Selaginella genomics database (<http://xselaginella.genomics.purdue.edu/>), and ChromDB (<http://www.chromdb.org/>) (Gendler et al., 2008) or through literature searches.

For all datasets, amino acid sequences were initially aligned using Clustal W and then adjusted by hand using MacVector (Cary, North Carolina). Maximum likelihood analysis was completed using RAxML (Stamatakis et al., 2008) as implemented by the CIPRES Science Gateway (<http://www.phylo.org/portal2/login!input.action>) (Miller et al., 2010). The model of amino acid evolution used was the default JTT. Bootstrap values are

presented at all nodes with greater than 50% support while nodes with less than 50% support are collapsed.

2.3 Quantitative real-time PCR

To assess expression of the PRC2 genes and VEL PHD family throughout the life cycle of *A. vulgaris*, the following tissue was collected from *A. vulgaris* plants: whole seedlings at the cotyledon, 1-3 leaf, and 6-8 leaf stages; leaves from 8-12 leaf stage plants; 8-12 leaf stage meristems (before vernalization); meristems subjected to 4 weeks of cold treatment at 4°C (during vernalization); meristems subjected to 8 weeks of cold treatment then removed to 18°C (after vernalization); inflorescence meristems; anthesis stage sepals, stamens and carpels; and developing fruits. At each stage, samples from three to ten different plants were collected and pooled. Total RNA was extracted using the RNeasy Mini kit (Qiagen, Valencia, CA). The RNA was treated with Turbo DNase (Ambion, Austin, TX) and cDNA was synthesized from 10 µg of total RNA using Superscript II reverse transcriptase (Invitrogen, Carlsbad, CA) and oligo (dT) primers.

Quantitative Real Time PCR (qRT-PCR) reactions were carried out using PerfeCTa SYBR Green FastMix Low Rox (Quanta Biosciences, Gaithersburg, MD) and analyzed in the Stratagene Mx3005P QPCR System (Agilent Technologies, Santa Clara, CA). Each 20 µl reaction included 4 µl of cDNA that had been diluted 1:5 and had a final primer concentration of 0.25 nmol/µL. A list of primers is included in Supplemental Table 1. Standard curves were run for all primer pairs to ensure high efficiency. The annealing temperature of all genes was 60° C with a 30 second extension. For each data point,

three technical replicates were analyzed. *AqIPP2* (isopentyl pyrophosphate:dimethylallyl pyrophosphate isomerase) expression was used for normalization.

2.4 Assessment of PRC2 Homolog Imprinting in *Aquilegia* Endosperm

In order to determine if any members of the *Aquilegia* PRC2 complex are imprinted in the endosperm, we took advantage of genetic polymorphisms between interfertile species of *Aquilegia*. We obtained several individual plants of *A. canadensis* and *A. vulgaris*. Total RNA was extracted from young leaves using the RNeasy kit (Qiagen, Valencia, CA) and cDNA was synthesized from 5ug RNA using Superscript II reverse transcriptase (Invitrogen, Carlsbad, CA). The 3' UTR of *AqSWN* and *AqCLF* were amplified by RT-PCR using Platinum Taq (Invitrogen, Carlsbad, CA) with specific primers (Supplemental Table 1) and purified using the QIAEX II Gel Extraction Kit (Qiagen, Valencia, CA) followed by column purification using the PCR Purification Kit (Qiagen, Valencia, CA). These gene fragments were then directly sequenced using Big Dye v3.1 (Life Technologies Corporation, Carlsbad, CA) and sequences were aligned with Clustal W using MacVector (Cary, North Carolina). This allowed the identification of single nucleotide polymorphisms (SNPs) representing restriction polymorphisms that could distinguish one of the *A. vulgaris* plants from one of the *A. canadensis* plants.

Several flowers on each of these plants were emasculated and reciprocal crosses were preformed. Seeds were collected when the seed coat was dark green and the endosperm had just cellularized (approximately a week after fertilization). At this stage, the *Aquilegia* embryo is approximately 1 mm in length out of a total seed length of 4 mm

and is tightly positioned at the micropylar end of the seed. Seeds were bisected horizontally to separate the embryo containing half from the endosperm-only half (Fig. 7A) and these separate samples were pooled to obtain 100mg of material for each.

RNA was extracted from the seeds using the method described by Vicient and Delseny (1999) with some modifications. The RNA was only extracted once in Phenol and the phenol:chloroform:isoamyl alcohol and chloroform:isoamyl alcohol steps were eliminated. The aqueous phase was then collected and separated into two 1.6 ml microcentrifuge tubes. A 0.1x volume of 3M sodium acetate and a 1.5x volume of ethanol was added to each tube and the mixture was stored overnight at -20° C. The tubes were then spun at 13,000 RPM for 30min at 4° C and the pellet was resuspended in 200µl of Lysis/Binding Solution from the RNAqueous kit (Ambion, Austin, TX) which was then used to further purify the RNA. RNA was treated with Turbo DNase (Ambion, Austin, TX).

cDNA was prepared from this RNA and the 3' end of *AqSWN* and *AqCLF* were amplified using the same methods as described above for parental leaves. For each digest, several RT-PCR amplifications were pooled before purification in order to obtain an adequately concentrated sample. *AqSWN* gene fragments from both seed halves and parental leaves were digested with Bpu10I in Buffer 3 (New England BioLabs, Ipswich, MA) for 2 hours at 37° C and run on a 2% agarose gel and visualized with Ethidium Bromide. *AqCLF* fragments from seeds and maternal leaves (control) were digested with AclI in Buffer 4 and 40 µM S-adenosylmethionine (New England BioLabs,

Ipswich, MA) for 16 hours at 37° C and run on a 2% agarose gel and visualized with Ethidium Bromide.

Section 3: Results and Discussion

3.1 Homologs of the PRC2 and the VEL PHD family in the *Aquilegia* Genome

We used a variety of bioinformatic approaches to identify PRC2 and VEL PHD homologs from the *Aquilegia coerulea* genome (http://www.phytozome.net/search.php?method=Org_Acoerulea). Similar to *Arabidopsis*, *A. coerulea* only has one ESC homolog, *AqFIE* (Fig. 1). Of the three identified homologs of *Multi Copy Suppressor of IRA (MSI)* (Fig. 2), one, *AqMSI1*, appears to be most similar to *MSI1* in *Arabidopsis thaliana*, which has been shown to associate with the PRC2 (De Lucia et al., 2008). The other two loci group with *A. thaliana MSI2* and *MSI3* (*AqMSI2*) or *A. thaliana FVE* and *MSI5* (*AqFVE*) (Hennig et al., 2003). While *Arabidopsis* has four homologs of Su(z)12 - *EMF2*, *VRN2*, *FIS2*, and *AT4G16810* (Chen et al., 2009) - we only identified one homolog in *A. coerulea*, *AqEMF2* (Fig. 3). This finding is consistent with other phylogenetic analyses of the VEFS domain containing proteins in plants which suggest that *VRN2*, *FIS2*, and *AT4G16810* are derived from rosid-specific duplication events, albeit with no statistical support (Chen et al., 2009). Note that although *AqEMF2* is not orthologous to *Arabidopsis EMF2*, we used the nomenclature of Cheng et al. (2009) in designating it *AqEMF2*. We recovered two E(z) homologs in *A. coerulea*, one that belongs to the *CLF* clade, *AqCLF*, and one from the *SWN* clade, *AqSWN* (Fig. 4). *A. thaliana* has three E(Z)-like genes, *CLF*, *SWN*, and *MEA* (Baumbusch et al., 2001), however, phylogenetic

analysis suggests that *MEA* is a product of a Brassicaceae-specific duplication of *SWN* (Spillane et al., 2007). We therefore conclude that relative to other model systems like rice and *Arabidopsis*, *Aquilegia* has a simpler complement of PRC2 homologs.

We also searched for homologs of the VEL PHD gene family, which include co-factors of PRC2 (Greb et al., 2007; De Lucia et al., 2008). These could not be identified from available annotated genes so we used a combination of DNA sequence similarity and gene prediction software

(<http://linux1.softberry.com/berry.phtml?topic=fgenes&group=programs&subgroup=gfind>) to identify four *A. coerulea* VEL PHD genes (Fig. 5). Our phylogenetic analysis shows that there are several clades within the angiosperm VEL PHD family. The first contains *A. thaliana* *VRN5* and one *A. coerulea* gene, *AqVRN5*. A second clade contains several *A. thaliana* genes including *VEL1*, *2*, and *3* and *VIN3* as well as two genes from *Aquilegia* termed *AqVIN3A* and *AqVIN3B*. A third clade contains one *A. coerulea* gene, *AqPHD1*, in addition to representatives from *Vitis* and rice but no apparent *Arabidopsis* homolog. This study indicates that while ancient duplications established these three main lineages, the *A. thaliana* gene family was strongly influenced by recent duplications that generated the four *VIN3/VEL1-3* loci.

In a further effort to annotate epigenetic loci from *A. coerulea*, other homologs of major gene lineages, including the PAF1 and SWR1 complexes as well as several genes thought to have PRC1-like function in plants, are shown in Supplemental Table 2.

These are purely bioinformatics identifications, however, unlike the PRC2 and VEL PHD homologs which were confirmed using RT-PCR.

3.2 Expression Analysis of the PRC2 and VEL PHD homologs in *A. vulgaris*

We characterized the expression of the five putative members of the PRC2 in *A. vulgaris* as well as the three VEL PHD genes most similar to the *A. thaliana* genes with known function. Tissue was collected at different stages throughout the life cycle of *Aquilegia vulgaris* and qRT-PCR was used to assess their expression. Three technical replicates were analyzed for each primer set on each sample and the data was normalized relative to the expression of the housekeeping gene *IPP2*.

We found that most PRC2 homologs are expressed at similar levels in all tissues and life stages sampled (Fig. 6A). One notable exception is *AqMSI1* whose expression level increases almost 10 fold in apical meristems during vernalization and almost 8 fold in early inflorescence meristems as compared to its expression at the cotyledon stage (Fig.6A). *MSI1* homologs in other species are known to participate in other chromatin remodeling complexes, including the CAF1 complex, so this increase in expression may be due to parallel functions and could reflect the large amount of chromatin remodeling necessary to complete these critical developmental transitions (Kohler et al., 2003). We also observed a small increase in *AqFIE* expression in both the fruits and the carpels, however, it is unclear if this increase in expression is functionally relevant because the expression levels of the other PRC2 members remain low in these tissues. We conclude that none of the PRC2 loci show stage or tissue-specific expression patterns,

consistent with the simple complement of PRC2 homologs in *Aquilegia* and the role they are hypothesized to play throughout development. In *A. thaliana*, the PRC2 gene *VRN2* similarly does not have a very dynamic expression pattern outside the developing seed, even during vernalization, despite the important role it plays at this stage (Gendall et al., 2001).

The VEL PHD finger family members, *AqVIN3A*, *AqVIN3B*, and *AqVRN5* are also expressed throughout *A. vulgaris* development (Fig.6B). *AqVIN3A* and *AqVRN5* peak in expression in the inflorescence and *AqVIN3B* expression is high at this stage as well. *AqVIN3B* expression peaks in the stamens while *AqVIN3A* is particularly low in this tissue. We cannot rule out that *A. vulgaris* VEL PHD proteins play a role in vernalization but, if they do, it does not appear to be mediated by specific expression patterns as with *VIN3* in *Arabidopsis* and wheat (Sung and Amasino, 2004; Fu et al., 2007). However, VEL PHD family members may also be involved in other aspects of plant development. For example, a rice VEL PHD gene, *LEAF INCLINATION 2*, has been shown to repress cell divisions in the region between the leaf blade and leaf sheath known as the collar and thus contribute to leaf angle, an important agricultural trait (Zhao et al., 2010). It may be interesting to further investigate the role of the VEL PHD genes in aspects of *Aquilegia* development beyond flowering time.

3.3 Parental Expression of *AqCLF* and *AqSWN* in *Aquilegia* Endosperm

As discussed above, select members of the PRC2 complex in several models exhibit parent-of-origin-specific patterns of imprinting and, hence, expression patterns in the

endosperm. We therefore sought to determine the imprinting patterns of specific PRC2 homologs in *Aquilegia*.

We chose to focus on the E(z) homologs *AqCLF* and *AqSWN* because in almost every case where PRC2 gene imprinting has been described, the targeted genes are one of several copies present in the genome, including the *SWN* paralogs *MEA* and *Mez1* in *Arabidopsis* and maize, respectively (Haun et al., 2007; Spillane et al., 2007; Rodrigues et al., 2010). Our experiment took advantage of the fact that many *Aquilegia* species are interfertile and their seeds have large, persistent endosperm (Fig. 7A) (Prazmo, 1965). Genetic variation between *Aquilegia* species is low and what variation exists is not fixed (Hodges and Arnold, 1994). Therefore we tested several plants and identified one *Aquilegia vulgaris* and one *Aquilegia canadensis* plant bearing polymorphisms in the 3'UTRs of *AqCLF* and *AqSWN* that could be distinguished by restriction digestion (Fig. 7B and C). We then conducted reciprocal crosses between the relevant individuals and collected the hybrid seeds. The seeds were bisected perpendicular to the micropyle to separate the endosperm from the embryo (Fig. 7A) and cDNA libraries were made from each half. We then amplified and digested the relevant 3' UTR fragments of *AqCLF* and *AqSWN* from both halves of the seeds as well as leaf tissue from both parents. We found no evidence for imprinting in either of these loci (Fig. 7B and C). While the parental alleles could be easily distinguished by restriction digest, the gene segments purified from the hybrid endosperm clearly contained both polymorphisms. While the *A. canadensis* allele of *AqSWN* appears to be present at a lower level than that of *A. vulgaris* in the endosperm sample from cross 1 (*A. canadensis* female x *A. vulgaris*

male), the alleles appear to be present at approximately equal levels in the endosperm sample from cross 2 (*A. vulgaris* female x *A. canadensis* male). This difference seems to be due to a stochastic variation in allele amplification since other duplicate reactions appeared equivalent, but we chose to show one entire set of concurrent reactions.

These findings do not mean that imprinting of other loci does not play a role in endosperm development in *Aquilegia*, on the contrary, it seems very likely that it does (Baroux et al. 2002). What they may suggest, however, is that the imprinting observed with Arabidopsis *MEA* and *FIS2*, as well as the grass loci *Mez1*, *ZmFIE1*, and *OsFIE1*, is related to the subfunctionalization of particular PRC2 paralogs for a role in endosperm development. In the cases of *MEA* and *FIS2*, this specialization is further associated with a higher rate of molecular evolution as indicated by statistical tests or exceedingly long-branch lengths (Spillane et al., 2007; Chen et al., 2009). Of course, this is not exclusively the rule as the single copy Arabidopsis locus *FIE* also shows imprinting (Ohad et al., 1996). Unfortunately, we were not able to identify suitable polymorphisms in *AqFIE* but, hopefully, such tests will be feasible in the future.

3.4 Conclusions

- *Aquilegia* has a simple complement of PRC2 homologs with no recent duplications
- The PRC2 genes are broadly expressed throughout *A. vulgaris* development with no obvious tissue or stage specialization

- The ancient paralogs, *AqCLF* and *AqSWN*, do not appear to be imprinted in *Aquilegia* endosperm.
- *Aquilegia* has four members of the VEL PHD family, three of which are similar to Arabidopsis genes known to function in flowering time.
- VEL PHD gene expression in *A. vulgaris* is not confined to vernalization as seen with *VIN3* in Arabidopsis, but is moderately increased both during vernalization and in the inflorescence.
- We have now identified a set of chromatin remodeling gene homologs in *Aquilegia* for further functional studies as well as phylogenetic analyses.

Acknowledgements

The authors would like to thank members of the Kramer lab and two anonymous reviewers for critical comments on the manuscript. This work was supported by NSF-BE award #0412727 to EMK (lead PI, Scott Hodges).

Supplementary data

Supplemental Tables 1 and 2

References

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J.,1990. BASIC LOCAL ALIGNMENT SEARCH TOOL. *Journal of Molecular Biology*. 215, 403-410.
- Ballerini, E.S. and Kramer, E.M.,2011. Environmental and molecular analysis of the floral transition in the lower eudicot *Aquilegia formosa*. *Evodevo*. 2, 4.

- Baroux, C., Spillane, C. and Grossniklaus, U.,2002. Evolutionary origins of the endosperm in flowering plants. *Genome Biology*. 3.
- Baumbusch, L.O., Thorstensen, T., Krauss, V., Fischer, A., Naumann, K., Assalkhou, R., Schulz, I., Reuter, G. and Aalen, R.B.,2001. The *Arabidopsis thaliana* genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes. *Nucleic Acids Research*. 29, 4319-4333.
- Beh, L.Y., Colwell, L.J. and Francis, N.J.,2012. A core subunit of Polycomb repressive complex 1 is broadly conserved in function but not primary sequence. *Proceedings of the National Academy of Sciences*. 109, E1063-E1071.
- Bratzel, F., López-Torrejón, G., Koch, M., Del Pozo, J.C. and Calonje, M.,2010. Keeping Cell Identity in *Arabidopsis* Requires PRC1 RING-Finger Homologs that Catalyze H2A Monoubiquitination. *Current Biology*. 20, 1853-1859.
- Calonje, M., Sanchez, R., Chen, L. and Sung, Z.R.,2008. EMBRYONIC FLOWER1 Participates in Polycomb Group–Mediated AG Gene Silencing in *Arabidopsis*. *The Plant Cell Online*. 20, 277-291.
- Chen, L.-J., Diao, Z.-Y., Specht, C. and Sung, Z.R.,2009. Molecular Evolution of VEF-Domain-Containing PcG Genes in Plants. *Molecular Plant*. 2, 738-754.
- De Lucia, F., Crevillen, P., Jones, A.M.E., Greb, T. and Dean, C.,2008. A PHD-Polycomb Repressive Complex 2 triggers the epigenetic silencing of FLC during vernalization. *Proceedings of the National Academy of Sciences*. 105, 16831-16836.

- Exner, V., Aichinger, E., Shu, H., Wildhaber, T., Alfarano, P., Caflisch, A., Gruissem, W., Köhler, C. and Hennig, L.,2009. The Chromodomain of LIKE HETEROCHROMATIN PROTEIN 1 Is Essential for H3K27me3 Binding and Function during Arabidopsis Development. PLoS ONE. 4, e5335.
- Feil, R.,2008. Epigenetics, an emerging discipline with broad implications. Comptes Rendus Biologies. 331, 837-843.
- Fu, D., Dunbar, M. and Dubcovsky, J.,2007. Wheat VIN3 -like PHD finger genes are up-regulated by vernalization. Molecular Genetics and Genomics. 277, 301-313.
- Gendall, A.R., Levy, Y.Y., Wilson, A. and Dean, C.,2001. The VERNALIZATION 2 Gene Mediates the Epigenetic Regulation of Vernalization in Arabidopsis. Cell. 107, 525-535.
- Gendler, K., Paulsen, T. and Napoli, C.,2008. ChromDB: The Chromatin Database. Nucleic Acids Research. 36, D298-D302.
- Goodrich, J., Puangsomlee, P., Martin, M., Long, D., Meyerowitz, E.M. and Coupland, G.,1997. A Polycomb-group gene regulates homeotic gene expression in Arabidopsis. Nature. 386, 44-51.
- Greb, T., Mylne, J.S., Crevillen, P., Geraldo, N., An, H., Gendall, A.R. and Dean, C.,2007. The PHD Finger Protein VRN5 Functions in the Epigenetic Silencing of Arabidopsis FLC. Current Biology. 17, 73-78.
- Guillon, A.-E., Page, D.R., Chambrier, P., Lionnet, C., Faure, J.-E., Grossniklaus, U. and Berger, F.,2004. Identification of new members of Fertilisation Independent Seed Polycomb Group pathway involved in the control of seed development in Arabidopsis thaliana. Development. 131, 2971-2981.

- Haun, W.J., Laouéillé-Duprat, S., O'Connell, M.J., Spillane, C., Grossniklaus, U., Phillips, A.R., Kaeppler, S.M. and Springer, N.M.,2007. Genomic imprinting, methylation and molecular evolution of maize Enhancer of zeste (Mez) homologs. *The Plant Journal*. 49, 325-337.
- Hennig, L., Taranto, P., Walser, M., Schönrock, N. and Gruissem, W.,2003. Arabidopsis MSI1 is required for epigenetic maintenance of reproductive development. *Development*. 130, 2555-2565.
- Hodges, S.A. and Arnold, M.L.,1994. COLUMBINES - A geographically widespread species flock. *Proceedings of the National Academy of Sciences of the United States of America*. 91, 5129-5132.
- Hodges, S.A. and Kramer, E.M.,2007. Columbines. *Current Biology*. 17, R992-R994.
- Holliday, R.,1994. Epigenetics: An overview. *Developmental Genetics*. 15, 453-457.
- Kinoshita, T., Yadegari, R., Harada, J.J., Goldberg, R.B. and Fischer, R.L.,1999. Imprinting of the MEDEA Polycomb Gene in the Arabidopsis Endosperm. *The Plant Cell*. 11, 1945-1952.
- Köhler, C. and Hennig, L.,2010. Regulation of cell identity by plant Polycomb and trithorax group proteins. *Current Opinion in Genetics & Development*. 20, 541-547.
- Kohler, C., Henning, L., Bouveret, R., Gheyselinck, J., Grossniklaus, U. and Gruissem, W.,2003. Arabidopsis MSI1 is a component of the MEA/FIE Polycomb group complex and required for seed development. *EMBO J*. 22, 4804-14.
- Kramer, E.M.,2009. Aquilegia: A New Model for Plant Development, Ecology, and Evolution. *Annual Review of Plant Biology*. 60, 261-277.

- Kramer, E.M. and Hodges, S.A.,2010. *Aquilegia* as a model system for the evolution and ecology of petals. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 365, 477-490.
- Kramer, Elena M., Stilio, Verónica S.D. and Schlüter, Philipp M.,2003. Complex Patterns of Gene Duplication in the *APETALA3* and *PISTILLATA* Lineages of the *Ranunculaceae*. *International Journal of Plant Sciences*. 164, 1-11.
- Lewis, E.,1978. A gene complex controlling segmentation in *Drosophila*. *Nature*. 276, 565-70.
- Luo, M., Platten, D., Chaudhury, A., Peacock, W.J. and Dennis, E.S.,2009. Expression, Imprinting, and Evolution of Rice Homologs of the Polycomb Group Genes. *Mol Plant*. 2, 711-723.
- Margueron, R. and Reinberg, D.,2011. The Polycomb complex PRC2 and its mark in life. *Nature*. 469, 343-349.
- Meyerowitz, E.M.,2002. Plants Compared to Animals: The Broadest Comparative Study of Development. *Science*. 295, 1482-1485.
- Miller, M.A., Pfeiffer, W. and Schwartz, T.,2010. Creating the CIPRES Science Gateway for Inference of Large Phylogenetic Trees. *Proceedings of the Gateway Computing Environments Workshop (GCE)*. New Orleans, LA, 1-8.
- Mosquna, A., Katz, A., Decker, E.L., Rensing, S.A., Reski, R. and Ohad, N.,2009. Regulation of stem cell maintenance by the Polycomb protein FIE has been conserved during land plant evolution. *Development*. 136, 2433-2444.
- Nekrasov, M., Wild, B. and Muller, J.,2005. Nucleosome binding and histone methyltransferase activity of *Drosophila* PRC2. *EMBO Rep*. 6, 348-53.

- Ohad, N., Margossian, L., Hsu, Y.C., Williams, C., Repetti, P. and Fischer, R.L.,1996. A mutation that allows endosperm development without fertilization. *Proceedings of the National Academy of Sciences*. 93, 5319-5324.
- Okano, Y., Aono, N., Hiwatashi, Y., Murata, T., Nishiyama, T., Ishikawa, T., Kubo, M. and Hasebe, M.,2009. A polycomb repressive complex 2 gene regulates apogamy and gives evolutionary insights into early land plant evolution. *Proceedings of the National Academy of Sciences*. 106, 16321-16326.
- Oliver, S.N., Finnegan, E.J., Dennis, E.S., Peacock, W.J. and Trevaskis, B.,2009. Vernalization-induced flowering in cereals is associated with changes in histone methylation at the *VERNALIZATION1* gene. *Proceedings of the National Academy of Sciences*. 106, 8386-8391.
- Pien, S. and Grossniklaus, U.,2007. Polycomb group and trithorax group proteins in *Arabidopsis*. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression*. 1769, 375-382.
- Prazmo, W.,1965. Cytogenetic studies on the genus *Aquilegia*. IV. Fertility relationships among the *Aquilegia* species. *Acta Soc Bot Pol*. 34, 667-685.
- Rodrigues, J., Luo, M., Berger, F. and Koltunow, A.,2010. Polycomb group gene function in sexual and asexual seed development in angiosperms. *Sexual Plant Reproduction*. 23, 123-133.
- Russo, V.E.A., Martienssen, R.A. and Riggs, A.D., *Epigenetic Mechanisms of Gene Regulation*. Cold Spring Harbor Laboratory Press, Woodbury, USA (1996).
- Schubert, D., Primavesi, L., Bishopp, A., Roberts, G., Doonan, J., Jenuwein, T. and Goodrich, J.,2006. Silencing by plant Polycomb-group genes requires dispersed

- trimethylation of histone H3 at lysine 27. EMBO J. 25, 4638-49. Epub 2006 Sep 7.
- Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B. and Cavalli, G.,2007. Genome Regulation by Polycomb and Trithorax Proteins. Cell. 128, 735-745.
- Spillane, C., Schmid, K.J., Laouelle-Duprat, S., Pien, S., Escobar-Restrepo, J.-M., Baroux, C.I., Gagliardini, V., Page, D.R., Wolfe, K.H. and Grossniklaus, U.,2007. Positive darwinian selection at the imprinted MEDEA locus in plants. Nature. 448, 349-52.
- Springer, N.M., Danilevskaya, O.N., Hermon, P., Helentjaris, T.G., Phillips, R.L., Kaeppeler, H.F. and Kaeppeler, S.M.,2002. Sequence Relationships, Conserved Domains, and Expression Patterns for Maize Homologs of the Polycomb Group Genes E(z), esc, and E(Pc). Plant Physiol. 128, 1332-1345.
- Stamatakis, A., Hoover, P. and Rougemont, J.,2008. A Rapid Bootstrap Algorithm for the RAxML Web Servers. Systematic Biology. 57, 758-771.
- Sung, S. and Amasino, R.M.,2004. Vernalization in Arabidopsis thaliana is mediated by the PHD finger protein VIN3. Nature. 427, 159-164.
- Sung, S., Schmitz, R.J. and Amasino, R.M.,2006. A PHD finger protein involved in both the vernalization and photoperiod pathways in Arabidopsis. Genes & Development. 20, 3244-3248.
- Vicient, C.M. and Delseny, M.,1999. Isolation of Total RNA from Arabidopsis thaliana Seeds. Analytical Biochemistry. 268, 412-413.
- Whitcomb, S.J., Basu, A., Allis, C.D. and Bernstein, E.,2007. Polycomb Group proteins: an evolutionary perspective. Trends in Genetics. 23, 494-502.

- Xu, L. and Shen, W.-H.,2008. Polycomb Silencing of KNOX Genes Confines Shoot Stem Cell Niches in Arabidopsis. *Current Biology*. 18, 1966-1971.
- Yoshida, N., Yanai, Y., Chen, L., Kato, Y., Hiratsuka, J., Miwa, T., Sung, Z.R. and Takahashi, S.,2001. EMBRYONIC FLOWER2, a Novel Polycomb Group Protein Homolog, Mediates Shoot Development and Flowering in Arabidopsis. *Plant Cell*. 13, 2471-2481.
- Zhao, S.-Q., Hu, J., Guo, L.-B., Qian, Q. and Xue, H.-W.,2010. Rice leaf inclination2, a VIN3-like protein, regulates leaf angle through modulating cell division of the collar. *Cell Res*. 20, 935-947.

Web References

- DFCI Plant Gene Index <http://compbio.dfci.harvard.edu/tgi/plant.html> September 19th 2011
- Phytozome http://www.phytozome.net/search.php?method=Org_Acoerulea September 19th 2011
- Selaginella Genomics <http://xselaginella.genomics.purdue.edu/> September 19th 2011
- Softberry FGENESH
<http://linux1.softberry.com/berry.phtml?topic=fgenes&group=programs&subgroup=gfind> September 19th 2011

Figure Legends

Figure 1: Maximum likelihood (ML) analysis of Extra Sex Combs

(ESC)/*FERTILIZATION INDEPENDENT ENDOSPERM (FIE)* homologs with ML bootstrap values shown at the nodes. Nodes with <50 support are collapsed.

At=*Arabidopsis thaliana*; Aq=*Aquilegia*. Arabidopsis sequences are indicated with a dot; *Aquilegia*, with an arrow.

Figure 2: Maximum likelihood (ML) analysis of homologs of the yeast protein Multi Copy Suppressor of IRA (MSI1) with ML bootstrap values shown at the nodes. Nodes with <50 support are collapsed. At=*Arabidopsis thaliana*; Aq=*Aquilegia*. Arabidopsis sequences are indicated with a dot; *Aquilegia*, with an arrow.

Figure 3: Maximum likelihood (ML) analysis of VEFS box containing proteins including *VERNALIZATION 2 (VRN2)*, *EMBRYONIC FLOWER 2 (EMF2)*, and *FERTILIZATION INDEPENDENT SEED 2 (FIS2)* with ML bootstrap values shown at the nodes. Nodes with <50 support are collapsed. At=*Arabidopsis thaliana*; Aq=*Aquilegia*. Arabidopsis sequences are indicated with a dot; *Aquilegia*, with an arrow.

Figure 4: Maximum likelihood (ML) analysis of the SET domain containing Enhancer of Zeste E(z) homologs including *CURLY LEAF (CLF)*, *SWINGER (SWI)*, and *MEDEA (MEA)* with ML bootstrap values shown at the nodes. Nodes with <50 support are

collapsed. At=*Arabidopsis thaliana*; Aq=*Aquilegia*. Arabidopsis sequences are indicated with a dot; *Aquilegia*, with an arrow.

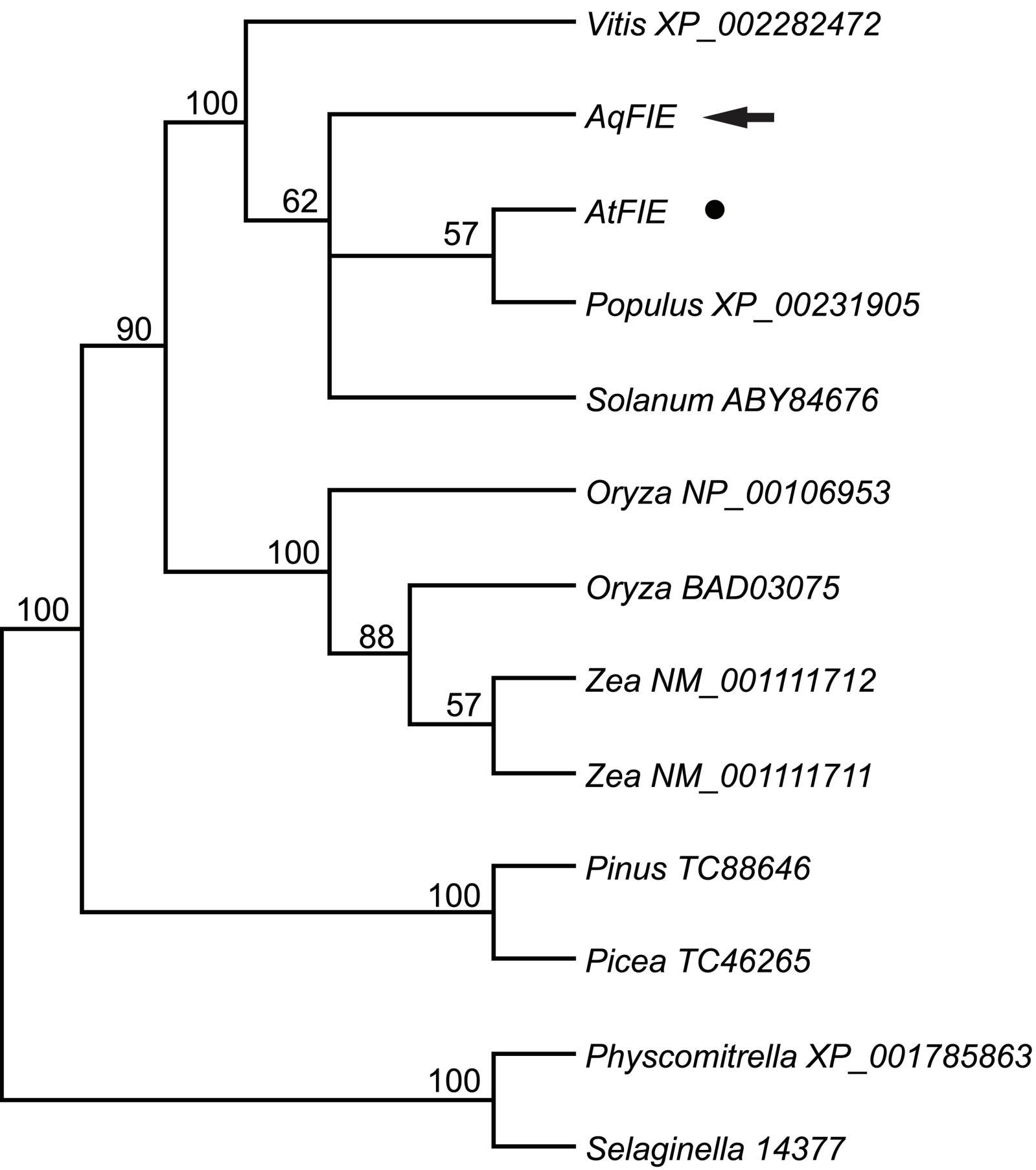
Figure 5: Maximum likelihood (ML) analysis of the VEL PHD family with ML bootstrap values shown at the nodes. Nodes with <50 support are collapsed. At=*Arabidopsis thaliana*; Aq=*Aquilegia*. Arabidopsis sequences are indicated with a dot; *Aquilegia*, with an arrow.

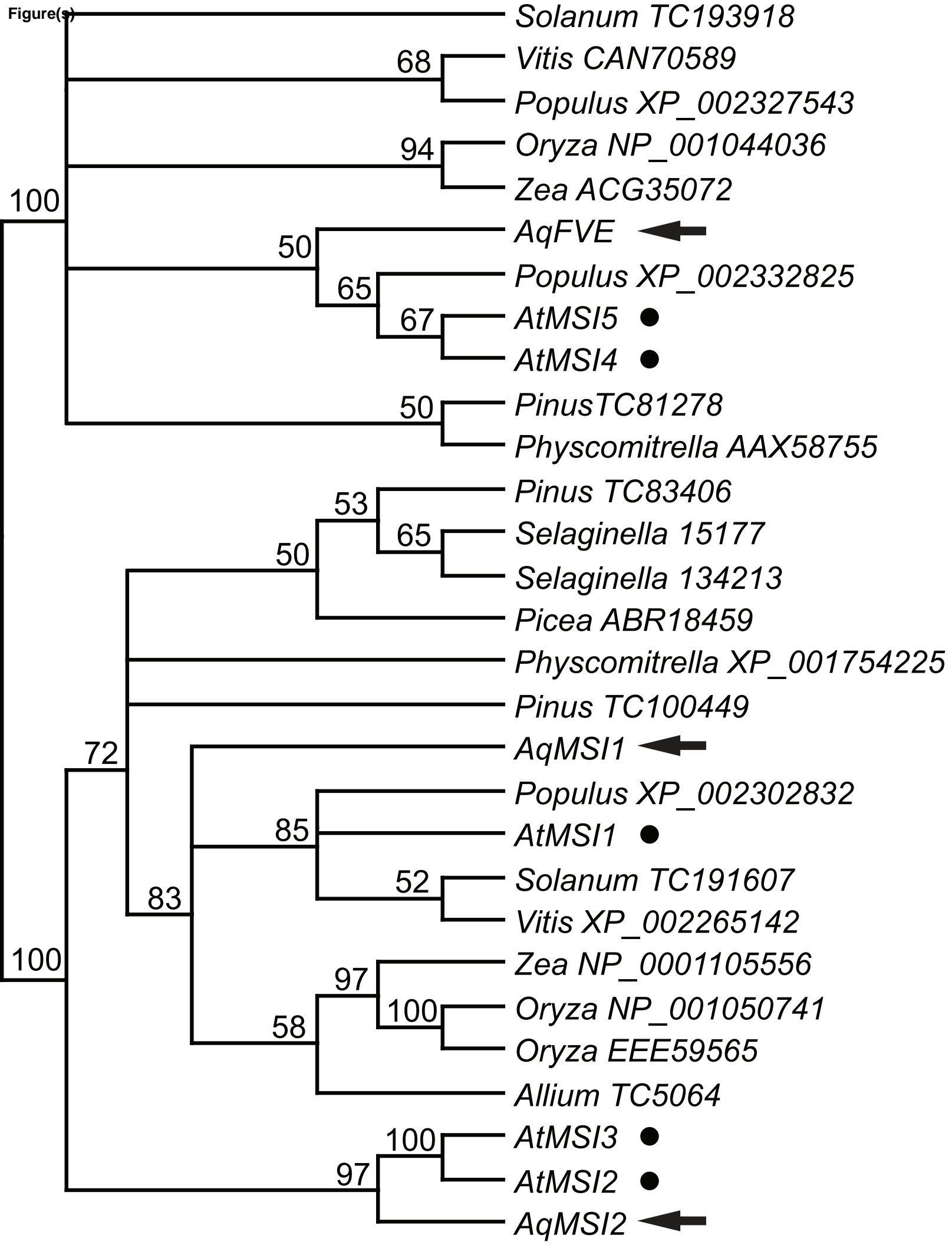
Figure 6: Quantitative Real Time PCR analysis of expression of the PRC2 complex and VEL PHD genes in *Aquilegia vulgaris*. Tissue from 3-10 plants was collected and pooled at each stage. For each data point, three technical replicates were analyzed. *Aq/PP2* expression was used for normalization. **A.** Average fold change of the PRC2 complex genes throughout the life cycle of *A. vulgaris* normalized to the cotyledon stage sample with SD error bars. **B.** Average fold change of the VEL PHD family throughout the life cycle of *A. vulgaris* normalized to the cotyledon stage sample with SD error bars.

Figure 7: *AqCLF* and *AqSWN* are not imprinted in *Aquilegia* endosperm. **A.** Diagram indicating position of embryo within the *Aquilegia* seed. Dotted line indicates orientation of seed bisection. **B.** *Acu I* digests of the 3' UTR of *AqCLF* purified from parental controls and hybrid seeds (both the endosperm (Ed) and embryo (Em) halves) run on a 2% agarose gel stained with ethidium bromide. *Acu I* cuts *Aquilegia canadensis* (Aqc) twice but only cuts *Aquilegia vulgaris* (Aqv) once. Hybrid seeds from reciprocal crosses (cross Em1/Ed1: *A.canadensis* female x *A. vulgaris* male, cross Em2/Ed2: *A. vulgaris*

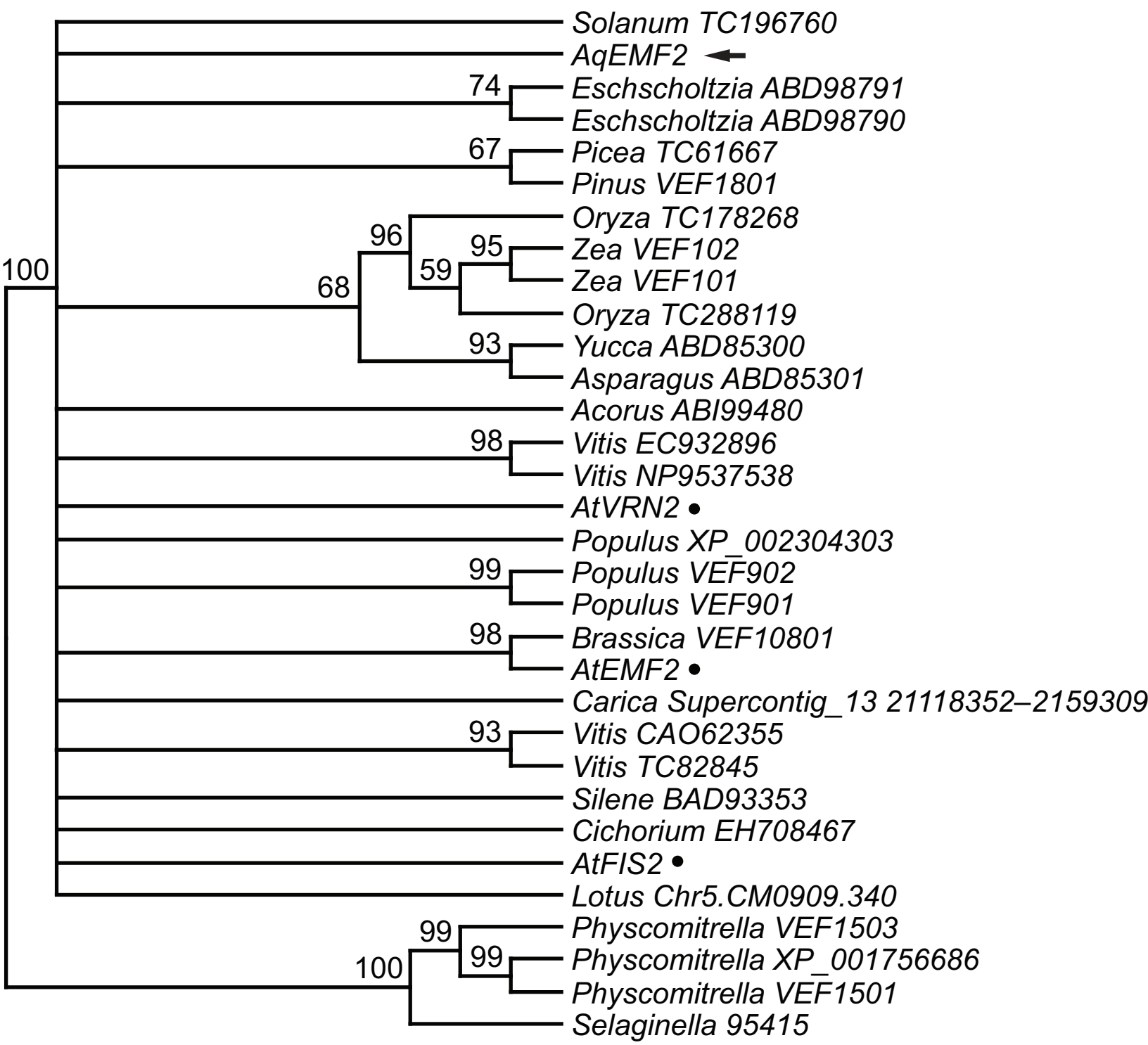
female x *A. canadensis* male) contain both polymorphisms. **B.** Bpu10I digests of the 3' UTR of *AqSWN* purified from parental controls and hybrid seeds (both the endosperm (Ed) and embryo (Em) halves) run on a 2% agarose gel stained with ethidium bromide. Bpu10I cuts *Aquilegia canadensis* (Aqc) once but does not cut *Aquilegia vulgaris* (Aqv). Hybrid seeds from reciprocal crosses (cross Em1/Ed1: *A. canadensis* female x *A. vulgaris* male, cross Em2/Ed2: *A. vulgaris* female x *A. canadensis* male) contain both polymorphisms.

Figure(s)

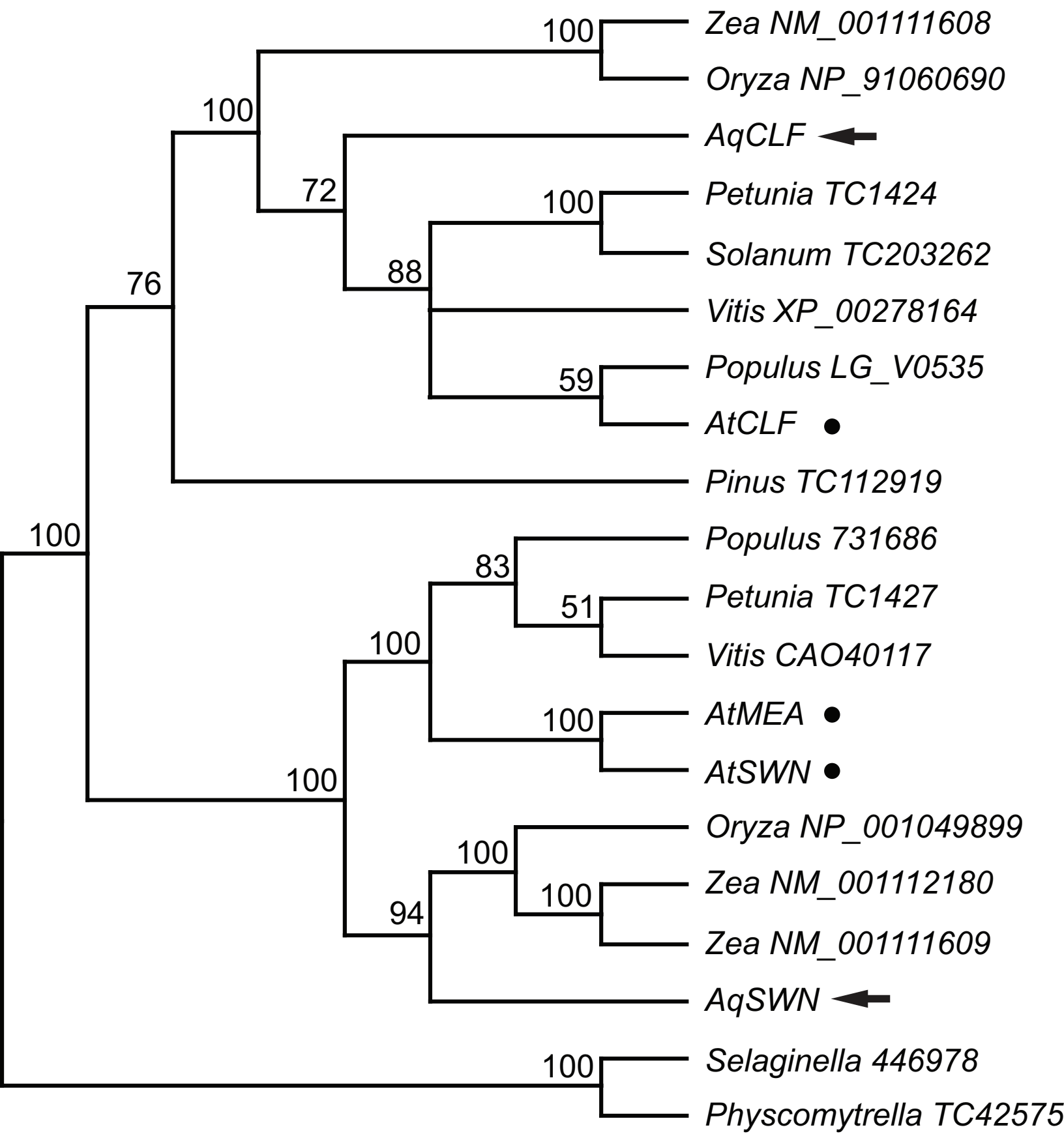




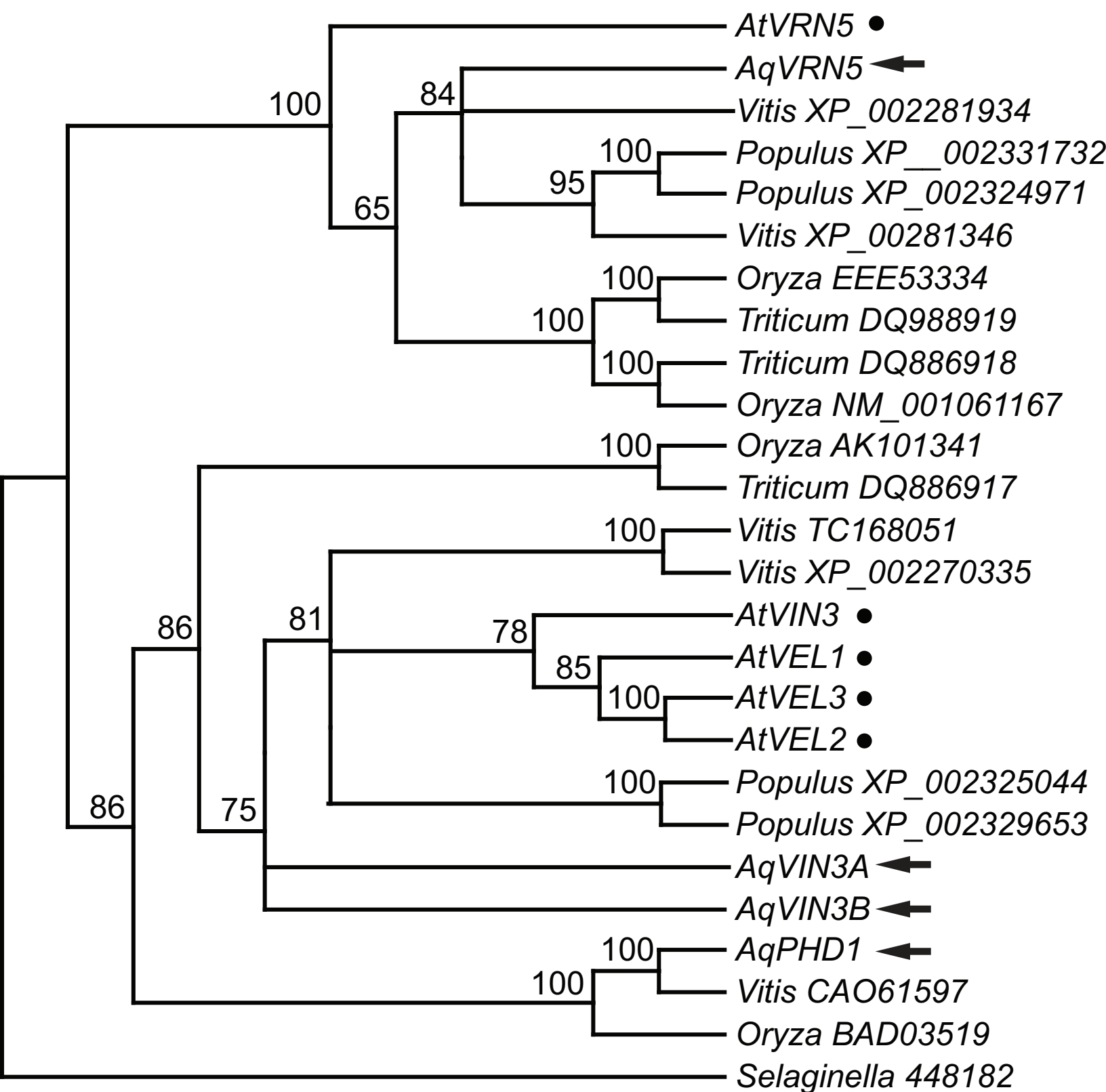
Figure(s)



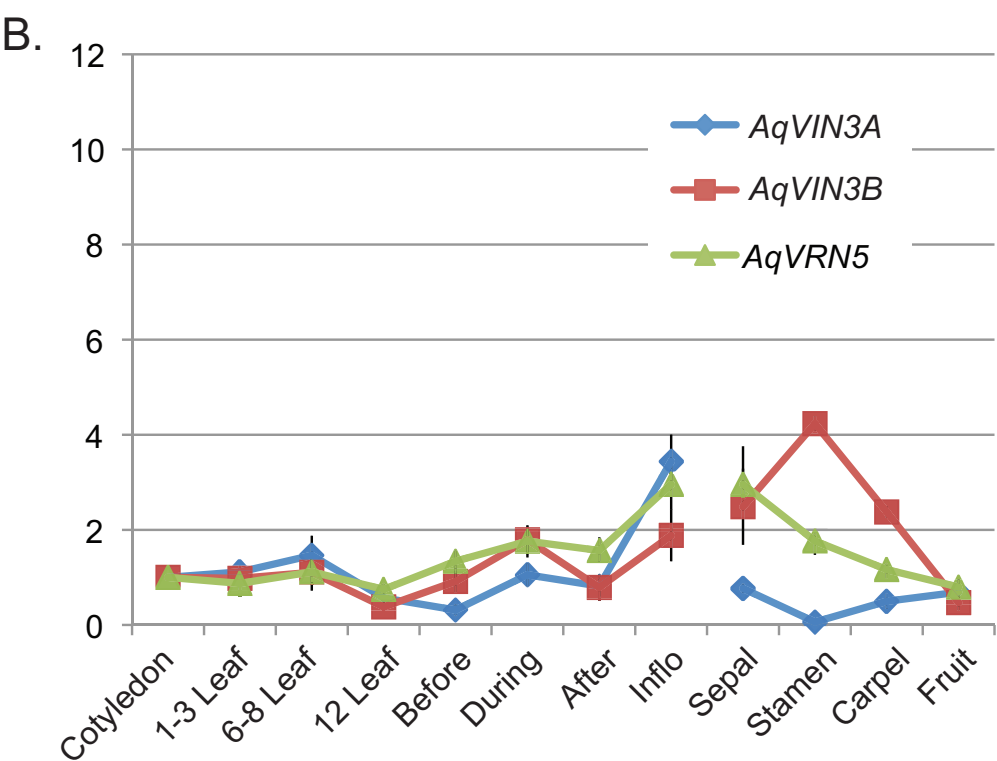
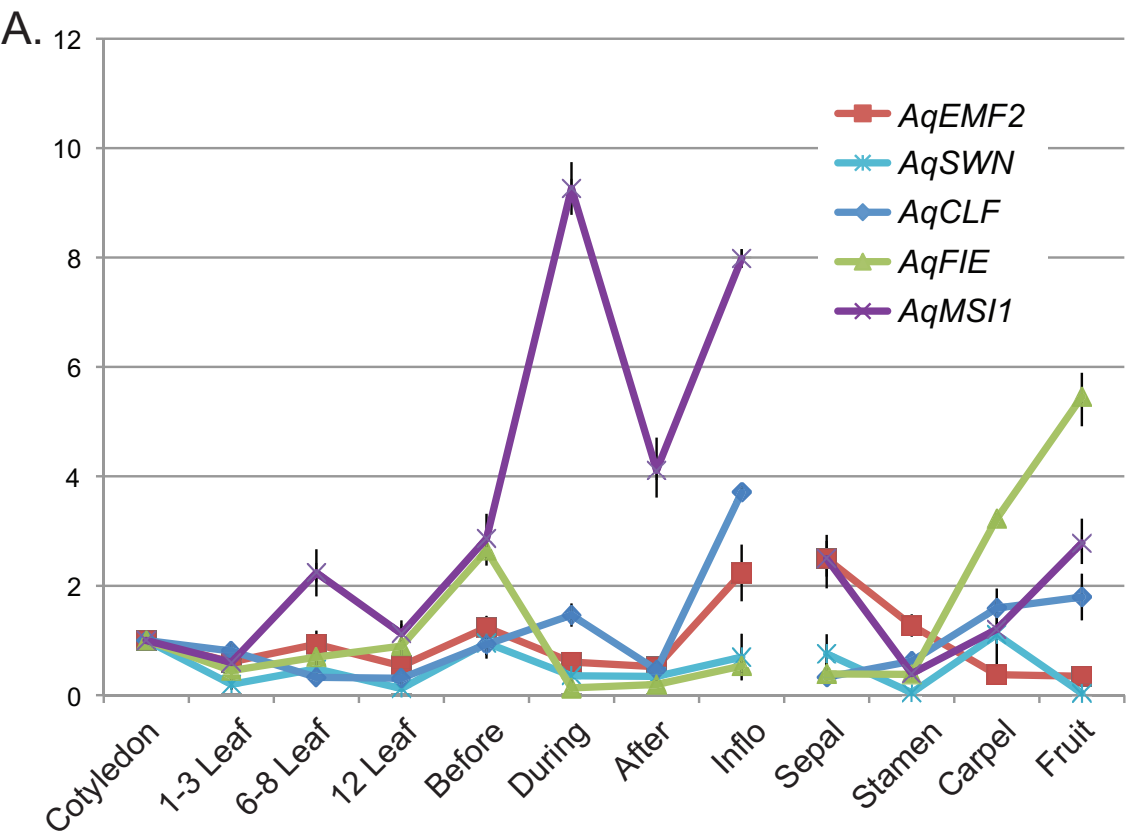
Figure(s)



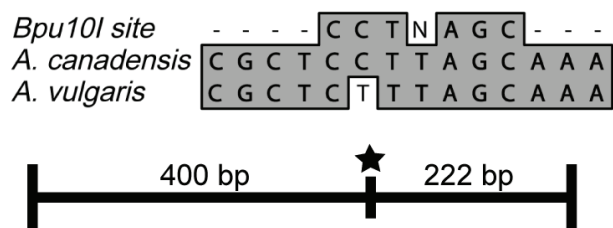
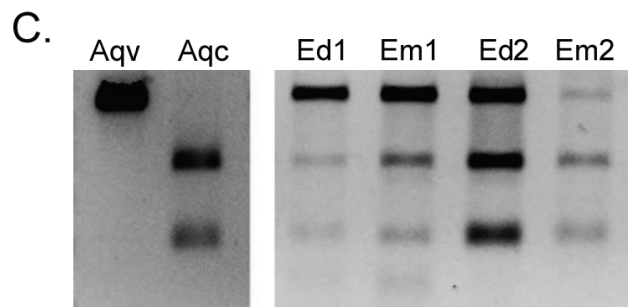
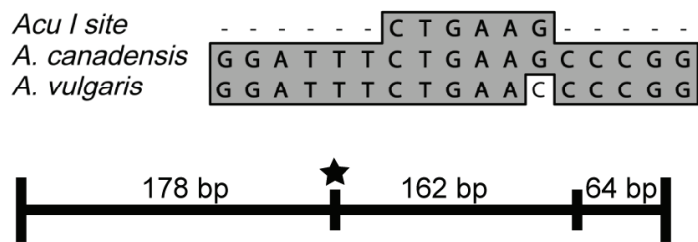
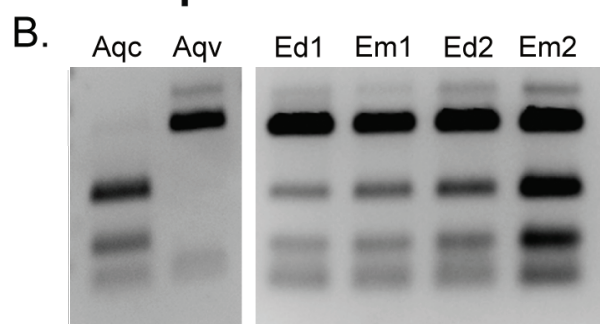
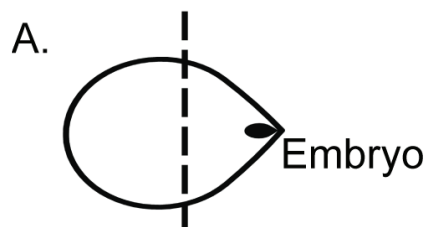
Figure(s)



Figure(s)



Figure(s)



Supplementary material for on-line publication only

[Click here to download Supplementary material for on-line publication only: Gleason and Kramer Suppl. Data.docx](#)